Article

Not peer-reviewed version

A Scale-Attention Optimized Hybrid Deep Learning Model for High-Accuracy Soil Salinity Mapping in Arid Oases

Mingjie Qian , Hangyuan Liu , Haoyi Wang , Shun Hu , Weitao Chen *

Posted Date: 17 April 2026

doi: 10.20944/preprints202604.1186.v1

Keywords: soil salinity; hyperspectral remote sensing; transform mechanism; deep learning

Article

A Scale-Attention Optimized Hybrid Deep Learning Model for High-Accuracy Soil Salinity Mapping in Arid Oases

Mingjie Qian [1,2], Hangyuan Liu [3], Haoyi Wang [4], Shun Hu [5] and Weitao Chen [4,*]

[1] School of Land Science and Technology, China University of Geosciences, Beijing 100083, China
[2] Key Laboratory of Land Consolidation and Rehabilitation, Ministry of Natural Resources, Beijing 100035, China
[3] Key Laboratory of Geological Survey and Evaluation of Ministry of Education, China University of Geosciences, Wuhan 430074, China
[4] School of Computer Science, China University of Geosciences, Wuhan 430074, China
[5] School of Environment Studies, China University of Geosciences, Wuhan 430074, China
* Correspondence: wtchen@cug.edu.cn

Abstract

Accurate monitoring of soil salinization in arid oasis regions is crucial for agricultural sustainability and ecological security. However, existing deep learning-based approaches often suffer from insufficient utilization of multi-scale information and inadequate modelling of feature interactions, limiting their accuracy in retrieving complex salinity patterns. To address these limitations, this study proposes a scale-attention optimized hybrid deep learning model that integrates multi-scale 1D convolutional neural networks (1D-CNN), bidirectional gated recurrent units (Bi-GRU), and Transformer mechanisms. The model employs a multi-scale feature extraction module to capture remote sensing signals across different scales, a scale attention mechanism to adaptively weight the most informative features, and a Bi-GRU-Transformer module to explore complex sequential and global feature relationships. The proposed framework is applied to the oasis irrigation zone in Weili County, Xinjiang, using hyperspectral data from the ZY-1E satellite, topographic indices, and spectral-derived variables. Experimental results demonstrate that our method achieves a coefficient of determination (R^2) of 0.952 and a root mean square error (RMSE) of 0.867 $g \cdot kg^{-1}$ on the test set, outperforming conventional 1D-CNN, GRU-Transformer, and other benchmark models with improvements of 2.8% in R^2 and 18.9% in RMSE.

Keywords: soil salinity; hyperspectral remote sensing; transform mechanism; deep learning

1. Introduction

Land degradation has emerged as a major global challenge for ecological protection and sustainable development [1]. Among various degradation processes, soil salinization is becoming increasingly severe, imposing significant pressure on agricultural production, ecological environments, and socio-economic development [2]. Consequently, accurate monitoring and retrieval of soil salinity have become focal points for researchers worldwide [3].

Traditional field-based methods for monitoring salt-affected soils, such as field investigations and soil sampling, are direct but time-consuming, labor-intensive, costly, and difficult to implement over large or inaccessible areas [4]. Remote sensing technology offers a promising alternative by enabling rapid and extensive identification and monitoring of soil salinity through the analysis of surface reflectance, terrain features, and land cover information [5]. In particular, the fusion of multi-source remote sensing data—such as hyperspectral imagery, digital elevation models (DEMs), and spectral indices—has shown great potential for improving the accuracy and reliability of land

 doi:10.20944/preprints202604.1186.v1

monitoring [6,7]. Hyperspectral data provide rich spectral information [8], DEM supply topographic details [9], and spectral indices capture essential indicators related to land cover and vegetation status [10]. However, integrating these multi-source datasets introduces complexity in data processing and analysis, posing a significant challenge for current remote sensing applications.

Accurate retrieval of soil salinity from multi-source remote sensing data has attracted considerable research attention. Spectral techniques, leveraging detailed hyperspectral information, have proven effective for rapid salinity assessment [11–13]. The reflectance of soil is influenced by its physical and chemical properties, with salt-affected soils often exhibiting distinct spectral absorption features that can be quantitatively related to salt content [14]. Salinization retrieval models have traditionally been constructed using either empirical statistical methods (e.g., linear and polynomial regression) or physical models. While empirical methods can achieve good accuracy under specific conditions [15], their generalization is limited by the representativeness of the sample data. Physically-based models are theoretically robust but often hindered by complex parameterization and high computational costs. Moreover, traditional approaches generally exhibit limited accuracy and fail to fully exploit the information potential of large-scale, multi-source remote sensing data [16–19].

In recent years, machine learning (ML) and deep learning (DL) techniques have significantly advanced the field of salinization retrieval [20]. Methods such as Random Forest (RF) [21], XGBOOST [22], Recurrent Neural Network (RNN) [23], Convolutional Neural Network (CNN) [24], and Transformer [25] have been widely applied. Compared to traditional techniques, DL models excel at processing and fusing large-scale, multi-source remote sensing data, capturing complex, nonlinear relationships between soil salinity and its influencing factors, thereby improving inversion accuracy and generalization performance. The powerful representation and automatic feature learning capabilities of DL enable the extraction of deep, nonlinear patterns from complex, high-dimensional data, leading to higher efficiency and precision [26,27].

As an important DL architecture, 1D-CNN has been widely used in salinity retrieval and environmental science due to its strength in extracting features from sequential data [28,29]. However, when dealing with large volumes of hyperspectral and multi-source data, 1D-CNN faces challenges such as overfitting and high computational costs [30]. Moreover, the diverse and intricate characteristics of salt-affected land may not be adequately captured by single-scale feature extraction, limiting the comprehensive representation of salinization features [31]. The Transformer architecture, capable of parallel processing across different spectral bands and data sources, can capture complexity within individual data sources and integrate information effectively [32,33]. Nevertheless, it has limitations in handling the inherent sequential properties of data and identifying deep-level correlations among complex features [34].

To address these challenges, this study proposes a novel salinization inversion method that integrates multi-scale feature extraction, a scale attention mechanism, and a Bi-GRU-Transformer module. The multi-scale feature extraction module captures feature information at different scales, providing a more comprehensive description of land salinization. The scale attention mechanism assigns adaptive weights to features at each scale, prioritizing the most informative ones. The Bi-GRU-Transformer mechanism combines the sequence-sensitive learning of Bi-GRU with the global correlation mining capability of Transformer, enhancing the model's ability to understand interactions among multiple features in the soil salinization process. This integrated framework aims to overcome the limitations of existing remote sensing techniques and improve the accuracy of salinized land identification and monitoring. The oasis irrigation area of Weili County, Xinjiang, is used as a case study to validate the performance of the proposed method.

2. Study Area and Available Data

2.1. Study Area

This study was conducted in the oasis irrigation area of Weili County, Xinjiang, China, to evaluate the performance of the proposed method. The area is located between 86°43′- 87°19′E and 40°44′—40°57′N. The soils are predominantly loam and sandy loam, and include saline, alkaline, and wind-blown sandy soils. Due to its inland location within the Eurasian continent, annual evaporation greatly exceeds precipitation, leading to widespread soil salinization.

2.2. Available Data and Preprocessing

The following types of data were collected and preprocessed for this study:

2.2.1. Soil Sampling Data

Field sampling was carried out in autumn 2021. A total of 132 sampling points were established, considering factors such as accessibility, geographical distribution, soil type, local cropping patterns, planting density, land cover, and land use to ensure data comprehensiveness and accuracy. Figure 1 shows the distribution of the study area and sampling points.

Given the significant spatial heterogeneity of soil moisture and salt content, a three-point composite sampling method was applied at each location. Soil samples were collected from the surface to a depth of 20 cm at three subsites around each sampling point, thoroughly mixed by quartering, and homogenized to represent the soil properties at that point. The geographic coordinates of each sampling point were recorded using GPS

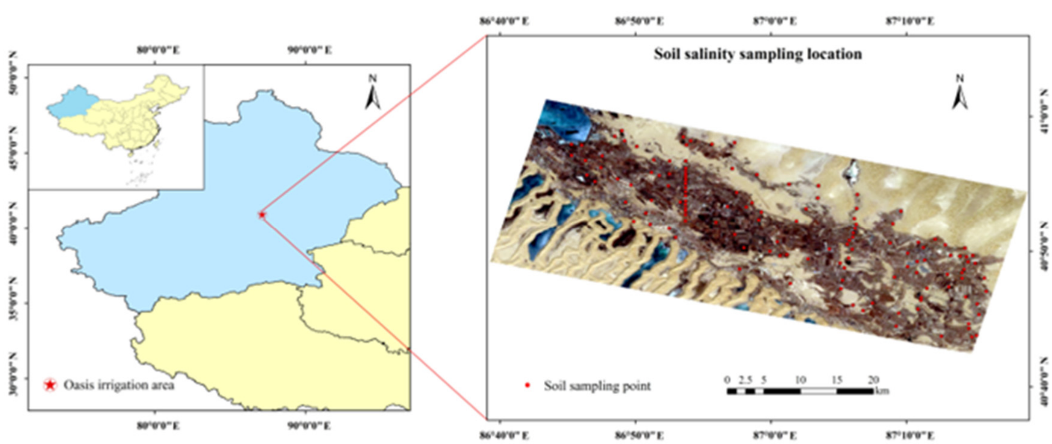


Figure 1. Distribution map of research area and sampling point.

Soil electrical conductivity (EC, μs·cm-1) was measured in the laboratory after extracting soil solutions with deionized water at a 1:5 soil-to-water ratio. Total soil salt content ($g \cdot kg^{-1}$) was estimated using the following calibration equation:

$$y = 0.0051x - 0.5241 \quad (1)$$

where x is the measured EC, y is the estimated salt content, and the coefficient of determination where the coefficient of determination R^2=0.9534.

2.2.2. Remote Sensing Data

Multiple remote sensing data sources were used, including spectral bands, spectral indices, and topographic data.

 doi:10.20944/preprints202604.1186.v1

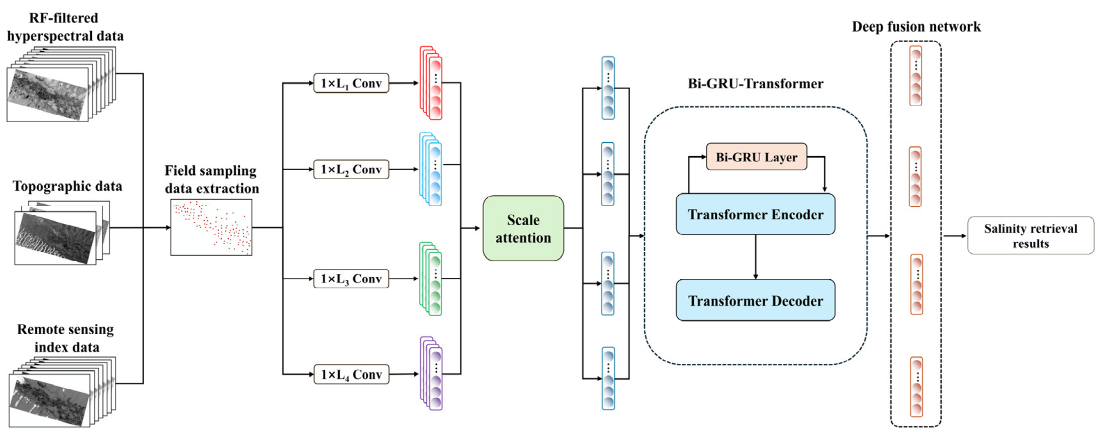
- Remote sensing data resources and spectral index data: Hyperspectral images were obtained from the Resource Satellite-1E (ZY-1E), covering the spectral range of 0.40–2.50 μm with a spectral resolution of 10 nm (VNIR) and 20 nm (SWIR), and a spatial resolution of 30 m. An image acquired on September 29, 2021, with minimal cloud cover, was selected. Using Recursive Feature Elimination with Cross-Validation [42] (Kumar et al. 2024), eight bands most representative for salinity map-ping were selected: three blue bands (B5, B6, B8), three green bands (B15, B16, B17), and two red bands (B25, B26). In addition, several spectral indices previously validated for salinity estimation were calculated; their formulas are listed in Table 1;
- Topographic data: To account for terrain effects on soil salinity, the Global Digital Elevation Model Ver-sion 3 (GDEMV3) was used to obtain elevation information at 30 m resolution. Slope and aspect were derived from the DEM. All topographic layers were clipped and georeferenced to match the study area;

Table 1. Spectral index calculation formula.

Salinity Index	Formulation	Reference	Description
NDSI	$(R-NIR)/(R+NIR)$	[35]	Reflects the content and distribution of salts in the soil
SI	$\sqrt{B\times R}$	[36]	Reflects soil salt content
SI1	$\sqrt{G\times R}$	[37]	Improved sensitivity and accuracy to salinity levels
SI2	$\sqrt{G^2+R^2+NIR^2}$	[38]	Improved sensitivity and accuracy to salinity levels
SI3	$\sqrt{G^2+R^2}$	[39]	Improved sensitivity and accuracy to salinity levels
SI4	$\sqrt{(B-G)^2+(G-R)^2}$	[40]	Improved sensitivity and accuracy to salinity levels
SI-T	$100(R-NIR)$	[41]	Improved accuracy of inversion of salinization extent

3. Methods

To address the challenge of accurately retrieving soil salinity from complex multi-source remote sensing data, this study proposes a novel method based on a scale-attention-optimized multi-scale 1D-CNN combined with a Bi-GRU-Transformer mechanism. As illustrated in Figure 2, the framework integrates multi-source inputs—including hyperspectral bands, spectral indices, and DEM-derived topographic features—to enrich feature representation. A multi-scale 1D-CNN serves as the backbone for feature extraction, enhanced by a scale-attention module that dynamically weights features across different scales. The refined features are then passed to a Bi-GRU-Transformer module, where Bi-GRU captures sequential dependencies and Transformer models global interactions among features. This combined design enables the model to better represent the nonlinear and multi-level relationships inherent in soil salinity data, thereby improving prediction accuracy and providing a more comprehensive understanding of salinization patterns.



Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 April 2026 doi:10.20944/preprints202604.1186.v1

Figure 2. Multiscale 1D-CNN and Dual Attention Mechanism Model Workflow Diagram.

3.1. Multi-Scale Feature Extraction Module

Conventional approaches often fail to fully exploit the rich scale-wise information contained in multi-source remote sensing data. To overcome this limitation, we developed a multi-scale feature extraction module based on a 1D-CNN architecture (Figure 3), aimed at capturing land salinization features across different spatial and spectral scales, thus enhancing the model's discriminative capacity.

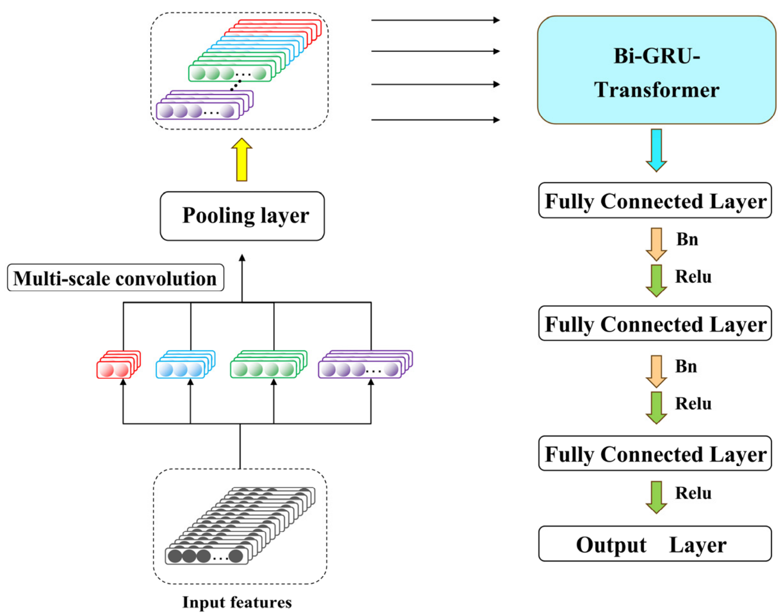


Figure 3. Multiscale Feature Extraction Module.

3.1.1. 1D-CNN Neural Network Model Structure

The designed 1D-CNN comprises the following key components:

- Convolutional Layers: These layers perform feature extraction and representation learning from the input data [43]. By employing convolutional kernels of varying sizes, the model progressively extracts feature at multiple scales, thereby strengthening its representational power;
- Pooling Layers: Pooling operations reduce the dimensionality of feature maps output by the convolutional layers, lowering computational cost and mitigating overfitting [44]. In this study, adaptive max pooling is applied, which automatically adjusts the pooling window size to retain the most salient features while improving generalization;
- Fully Connected (FC) Layers: FC layers transform the high-level features into final outputs for regression. Multiple FC layers are used to further abstract features, coupled with batch normalization (BN) and ReLU activation to enhance learning stability and expressive capability;

3.1.2. Multi-Scale Feature Extraction

The module employs multiple parallel 1D convolutional branches with different kernel sizes (denoted as L_1, L_2, …, L_n). Each branch processes the input multi-source remote sensing sequence with a set of kernels of a given scale:

$$d_l = Sigmoid(k_l \times x_l + b_l) \quad (2)$$

where x_l denotes the input feature vector, k_l the convolutional kernel with scale l, and d_l is the final output obtained by convolving the input hyperspectral data sequence of salinization from

 doi:10.20944/preprints202604.1186.v1

multiple sources. To align the lengths of outputs from different branches, scale-specific pooling is applied after each convolution. The pooled multi-scale features are then forwarded to the Bi-GRU-Transformer module. A deep fusion network—consisting of an input layer, three FC layers with BN and ReLU, and an output layer—is employed for final regression. The model is trained by minimizing the mean squared error (MSE)loss:

$$L = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \quad (3)$$

where N is the number of samples, y_i the true salinity value, and $\hat{y}_i$ the predicted value. Parameters are updated via gradient descent.

3.2. Scale Attention Mechanism

Features extracted by the multi-scale 1D-CNN originate from different spectral bands and receptive fields, and their contributions to the final salinity prediction are not equal. Direct concatenation may dilute the influence of more discriminative scales. To address this, a scale attention mechanism is introduced to adaptively weight and recombine multi-scale features based on their importance, thereby preserving the most relevant information for the regression task.

As shown in Figure 4, the feature maps F_i from the i –th scale first undergoes a linear transformation:

$$\hat{Z}_i = W_i \cdot F_i + b_i \quad (4)$$

where W_i and b_i are learnable parameters initialized using the He strategy. An attention score is then computed via softmax normalization:

$$A_i = softmax(\hat{Z}_i) = e^{\hat{Z}_i} / \sum_{j=1}^{N} e^{\hat{Z}_i} \quad (5)$$

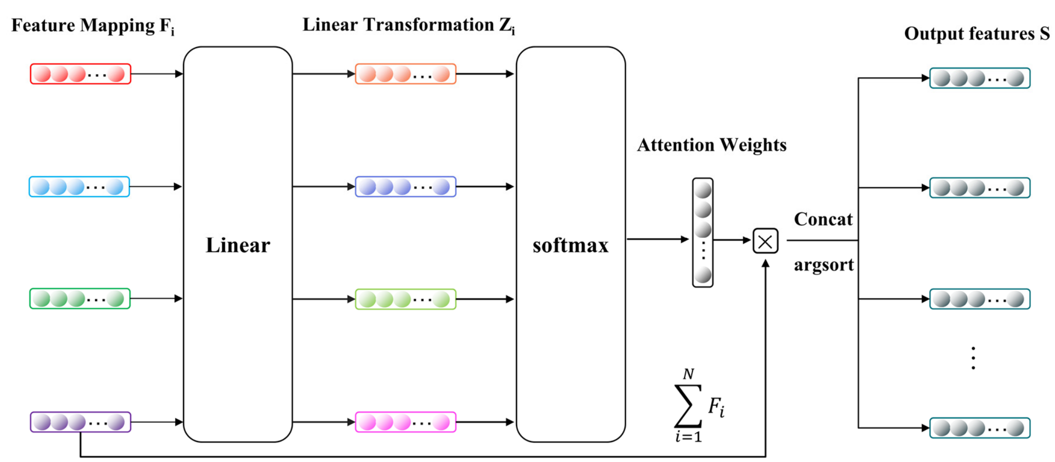
A_i represents the attention score for the i th scale feature, and $e^{\hat{Z}_i}$ ensures that the score is non-negative. The softmax layer transforms the output of the linear layer mapping into a probability distribution, representing the weights of different features in the model. Once the attention weights are determined, the feature mappings are reweighted. Based on the attention weights, the original feature mappings are reordered and recombined:

$$F_i' = A_i \cdot F_i \quad (6)$$

$$Index_{sorted} = agsort(A_i) \quad (7)$$

$$S = Concat\left(F'_{index(1)}, \dots, F'_{index(n)}\right) \quad (8)$$

where F_i' is the weighted feature mapping, the $agsort$ function returns the indices of feature importance according to the attention weights, and $Index_{sorted}$ is the feature set sorted according to the attention weights. All weighted features are combined to obtain the weighted feature S.



Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 April 2026 doi:10.20944/preprints202604.1186.v1

Figure 4. Scale Attention.

3.3. Bi-GRU-Transformer Mechanism

To effectively capture both sequential dependencies and global feature interactions in salinization data, we propose a hybrid Bi-GRU-Transformer module (Figure 5). This design integrates the sequential modeling strength of Bi-GRU with the global relational capacity of Transformer, offering a more comprehensive feature analysis framework.

The Transformer component follows the standard encoder architecture [45], employing multi-head self-attention to model interactions across all input positions:

$$Attention(Q,K,V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (9)$$

$$Q = xW^Q, K = xW^K, V = xW^V \quad (10)$$

$$head_i = Attention\left(QW_i^{\,Q}, KW_i^{\,K}, VW_i^{\,V}\right) \quad (11)$$

$$Multihead(Q,K,V) = Concat(head_1, \ldots, head_h)W^O \quad (12)$$

where Q, K, and V are the query, key, and value matrices, respectively, and d_k is the dimensionality of the keys, and h is the number of heads.

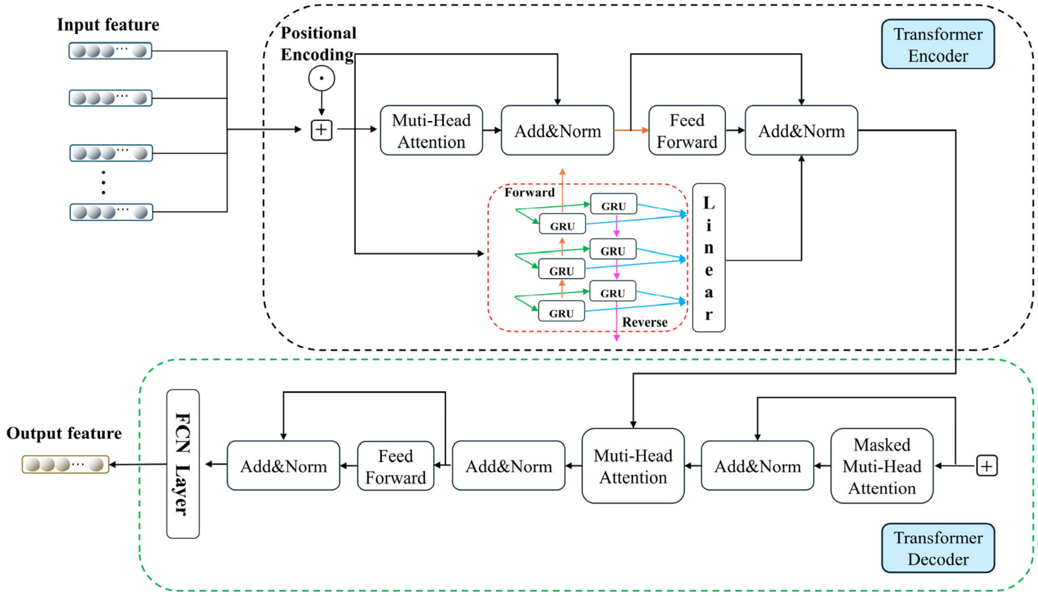
In parallel, a Bi-GRU processes the same input sequence to capture bidirectional contextual information:

$$\overrightarrow{h_z} = \overrightarrow{GRU}(\overrightarrow{x_z}) \quad (13)$$

$$\overleftarrow{h_z} = \overleftarrow{GRU}(\overleftarrow{x_z}) \quad (14)$$

$$h_z = \overrightarrow{h_z} \oplus (\overleftarrow{h_z}) \quad (15)$$

where, $\overrightarrow{h_z}$ and $\overleftarrow{h_z}$ represent the hidden states of the forward and backward GRU, respectively, at the importance ranking z, and $\oplus$ denotes the concatenation operation. The outputs from the Transformer encoder and Bi-GRU are concatenated and passed to subsequent layers. Crucially, before entering this module, features are sorted by the scale attention scores, ensuring that the model prioritizes the most salient information. This combined mechanism enables the model to simultaneously learn local sequential patterns and global feature relationships, significantly improving its ability to interpret complex soil salinity data.



 doi:10.20944/preprints202604.1186.v1

Figure 5. Bi-GRU-Transformer Mechanism.

4. Experimental Setup

4.1. Dataset

The experimental dataset comprised 18 predictive features derived from multi-source remote sensing data, including hyperspectral bands, spectral indices, and topographic factors. Each of the 132 sampling points was associated with this 18-dimensional feature vector. The dataset was randomly split into a training set and a test set at a ratio of 7:3 for model development and evaluation.

4.2. Neural Network Parameter Settings

The network architecture and hyper parameters were configured as follows. The model was trained for 500 epochs using the Adam optimizer with an initial learning rate of 0.001 and a batch size of 64. The minimum learning rate was set to 1e-6. The multi-scale 1D-CNN component employed four convolutional kernels of different sizes: 1, 3, 5, and 7. The detailed structure of the backbone network is provided in Table 2. In the Bi-GRU-Transformer module, the hidden dimensions of the Transformer and Bi-GRU were set to 256 and 128, respectively; the Bi-GRU consisted of two layers, and the number of attention heads was set to 9.

Table 2. Backbone Network Structure.

Layer name	Number	Kernel size	Stride	Padding
$Conv_{L1}$	64	1×1	1×1	0
Max pooling	/	2×1	2×1	/
$Conv_{L2}$	64	3×1	1×1	1
Max pooling	/	2×1	2×1	/
$Conv_{L3}$	64	5×1	1×1	2
Max pooling	/	2×1	2×1	/
$Conv_{L4}$	64	7×1	1×1	3
Max pooling	/	2×1	2×1	/
Full Connected1	256	/	/	/
Full Connected2	128	/	/	/
Full Connected3	64	/	/	/

4.3. Evaluation Metrics

Model performance was assessed using four widely adopted accuracy metrics: the coefficient of determination (R^2), root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE) [46]. These metrics provide a comprehensive evaluation of the model's precision and stability in soil salinity estimation.

4.4. Model Comparison

To thoroughly evaluate the performance of the proposed fusion model, we compared it with several representative machine learning and deep learning benchmarks, including Random Forest (RF), standard 1D-CNN, 1D-CNN with multi-head attention, GRU-Transformer, and SSA-LSTM-Transformer [47–51]. These comparisons help elucidate the relative strengths and potential limitations of different modeling approaches when applied to complex environmental remote sensing data.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 April 2026 doi:10.20944/preprints202604.1186.v1

5. Results and Discussion

5.1. Analysis of Soil Salinity Sampling Data

The statistical summary of soil salinity data is presented in Table 3. Following the Xinjiang soil salinization classification (Mu et al., 2012), surface soils (0–20 cm) were categorized into five classes: non-saline (0–8 g/kg), slightly saline (8–10 g/kg), moderately saline (10–15 g/kg), heavily saline (15–20 g/kg), and saline (>20 g/kg). High coefficients of variation were observed across the total dataset (38.6%), training set (39.1%), and test set (38.2%), confirming significant spatial heterogeneity and supporting the need for a robust, scale-aware retrieval model.

Table 3. Summary Statistics of Soil Salinity.

Dataset	Sample size	Non-saline Soil	Slightly Saline Soil	Moderately Saline Soil	Heavily Saline Soil	Saline Soil	Soil salinity (g/kg)				std
							max	min	mean	median	
Total	132	46	40	32	11	3	26.07	2.05	9.78	8.47	3.78
Train set	92	31	28	23	8	2	26.07	2.05	9.82	8.47	3.84
Test set	40	14	13	9	3	1	23.94	3.97	9.69	8.44	3.70

5.2. Correlation between Multi-Source Remote Sensing Data and Soil Salinity

Pearson correlation analysis (Table 4) revealed that all 18 remote sensing predictors were significantly correlated ($p < 0.05$) with measured soil salinity. Reflectance from the eight selected hyperspectral bands showed moderate positive correlations ($|r| \approx 0.5$). Spectral indices, particularly NDSI ($r = 0.68$) and SI-T ($r = 0.62$), exhibited stronger associations, while topographic factors such as slope also showed meaningful relationships. These results validate the relevance of the integrated multi-source feature set for salinity modeling.

Table 4. Correlation Coefficient of Multi-Source Remote Sensing Data.

Hyperspectral	Correlation	Salinity index	Correlation	Terrain factor	Correlation
B5	0.49	NDSI	0.68	DEM	0.49
B6	0.46	SI	0.30	Aspect	0.15
B8	0.54	SI1	0.49	Slpoe	0.43
B15	0.49	SI2	0.46		
B16	0.50	.SI3	0.61		
B17	0.52	SI4	0.54		
B26	0.44	SI-T	0.62		
B27	0.43				

5.3. Model Evaluation

5.3.1. Algorithm Error and Accuracy

The proposed model demonstrated superior convergence and generalization (Figure 6), achieving the lowest steady-state loss on both training and test sets. As shown in Figure 7, it also attained the highest R^2 values, significantly outperforming the standard 1D-CNN and showing clear gains over attention-enhanced and sequence-based variants (e.g., 1D-CNN with multi-head attention, GRU-Transformer). An ablation on attention heads (Figures 8 and 9) indicated that 9 heads yielded optimal learning dynamics and stability, enabling effective multi-subspace representation.

 doi:10.20944/preprints202604.1186.v1

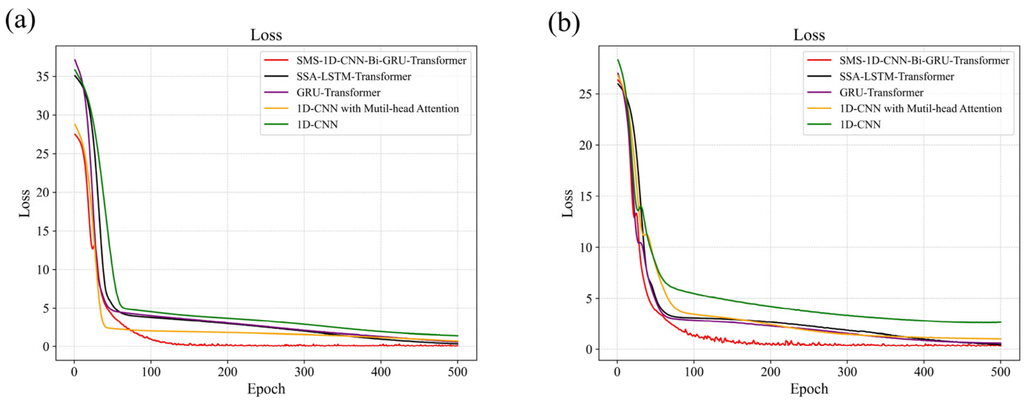


Figure 6. Loss Functions for Different Various Neural Networks: (a) Training Set(b) Test Set.

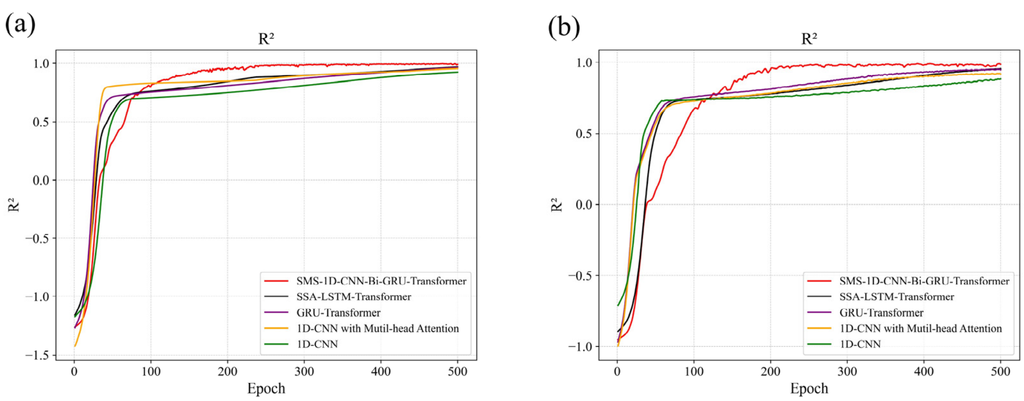


Figure 7. R^2 for Different Various Neural Networks: (a) Training Set (b) Test Set.

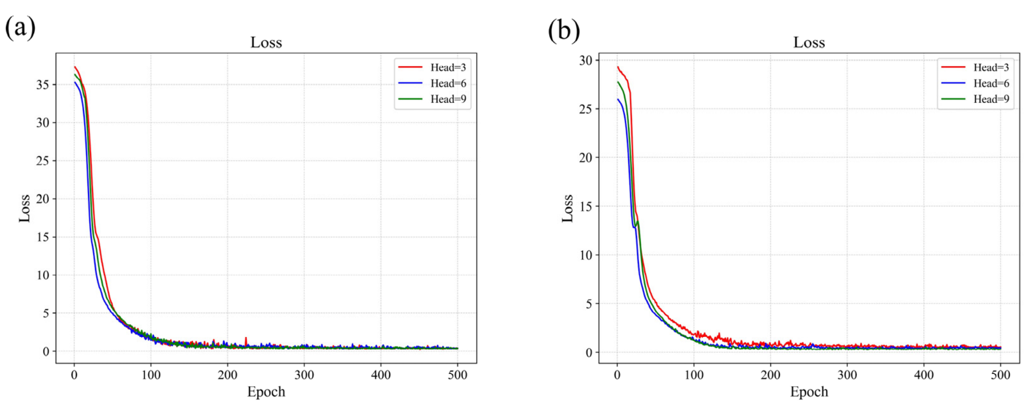


Figure 8. Loss Function for Different Attention Heads: (a) Training Set (b) Test Set.

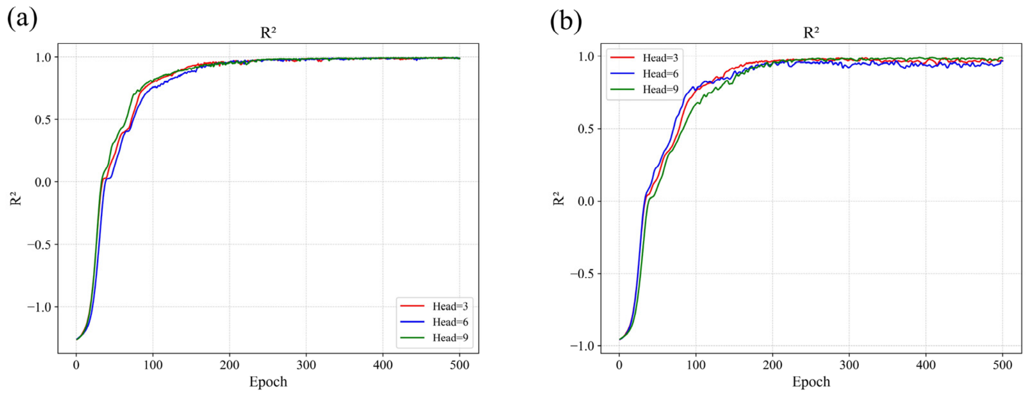


Figure 9. R^2 for Different Attention Heads: (a) Training Set (b) Test Set.

5.3.2. Regression Performance Evaluation

Quantitative results (Figure 10, Table 5) further substantiate the model's advantages. The proposed framework achieved an R^2 of 0.952, with an RMSE of 0.867 g·kg⁻¹, MAE of 0.451 g·kg⁻¹, and MAPE of 4.853%, outperforming all benchmark models. The integration of scale-aware feature weighting and hybrid sequence-global modeling is thus proven effective in reducing systematic prediction error and enhancing model reliability.

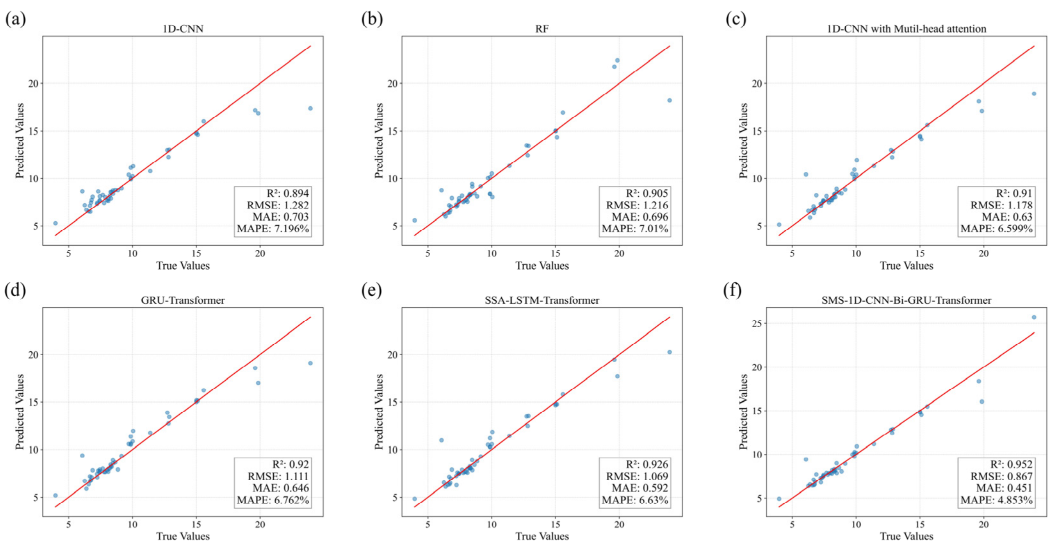


Figure 10. Accuracy Evaluation of Different Algorithms: (a)BP(b)1D-CNN(c)1D-CNN with self-attention(d)1D-CNN with Mutil-head attention(e)GRU-Transformer(f) MS-1D-CNN with Bi-GRU-Transformer Optimized by Scale attention.

Table 5. Accuracy Evaluation of Different Algorithms.

Methods	R^2	RMSE/(g·kg-1)	MAE/(g·kg-1)	MAPE/%
1D-CNN	0.894	1.282	0.703	7.196
RF	0.905	1.216	0.696	7.01
1D-CNN with Mutil-head attention	0.91	1.178	0.63	6.599
GRU-Transformer	0.92	1.111	0.646	6.762
SSA-LSTM-Transformer	0.926	1.069	0.592	6.63
SMS-1D-CNN-Bi-GRU-Transformer	**0.952**	**0.867**	**0.451**	**4.853**

5.4. Visual Interpretation of Retrieval Performance and Spatial Patterns

The spatial distribution of retrieved salinity classes (Figure 12) reveals a clear eco-topographic pattern: non-saline soils dominate well-drained higher terrains (67.70% of area), while saline classes progressively concentrate in low-lying depressions and inadequately drained irrigated zones. This alignment with known hydro-geomorphic controls underscores the model's physical plausibility.

A detailed visual comparison with benchmark models (Figure 13) highlights the superior categorical accuracy of the proposed approach. For example, in estimating non-saline areas—critical for agricultural planning—our model reduced the proportion error to only 3.7%, compared to deviations of 4.17–8.82% from other models. More notably, for spatially limited but agronomically critical classes (e.g., heavily saline and saline soils), our model maintained high fidelity, with errors lower by 2.06–5.91% and 5.49–17.42%, respectively. This indicates that the multi-scale attention mechanism and Bi-GRU-Transformer module jointly improve sensitivity to both dominant and rare but impactful salinity levels, enabling more reliable delineation of salinity hotspots for targeted management.

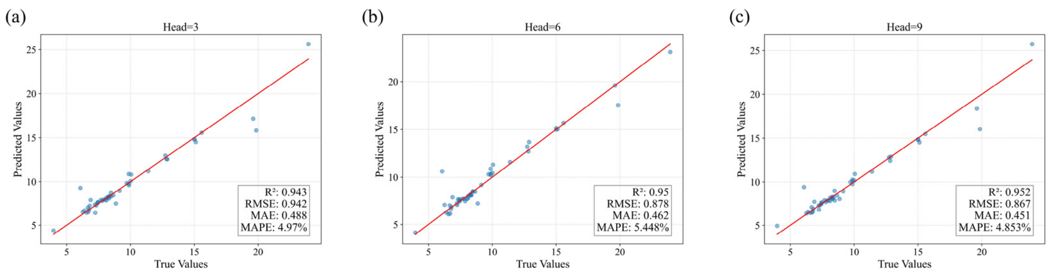


Figure 11. Accuracy Evaluation for Different Attention Heads: (a) Head=3(b) Head=6(c) Head=9.

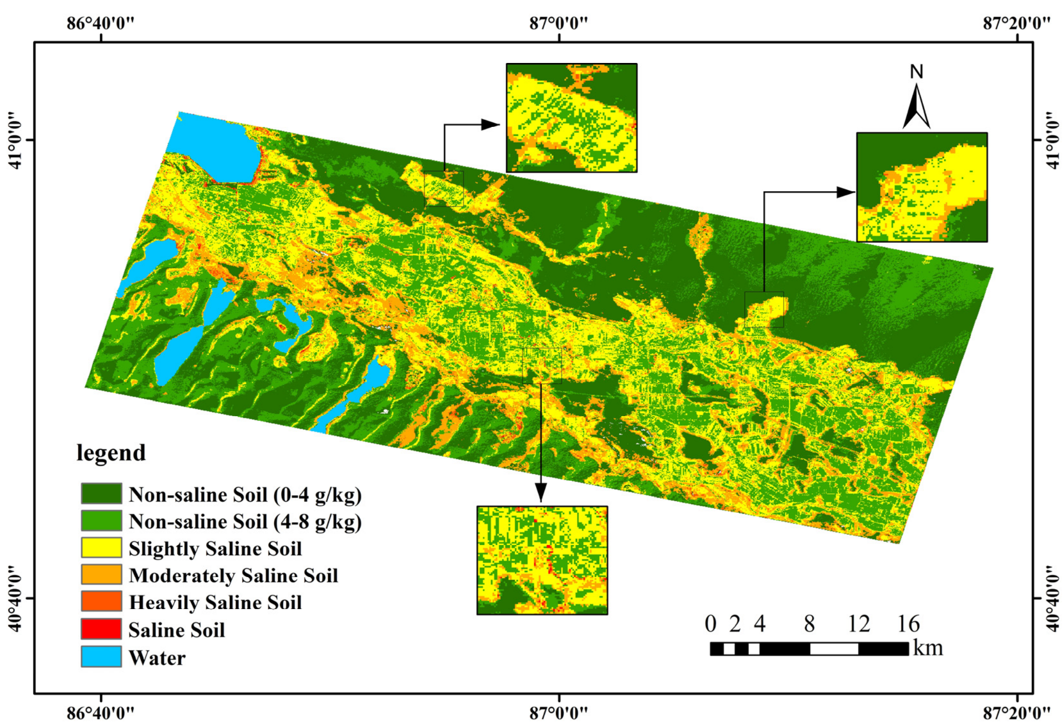


Figure 12. Soil Salinity Retrieval Result.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 April 2026 doi:10.20944/preprints202604.1186.v1

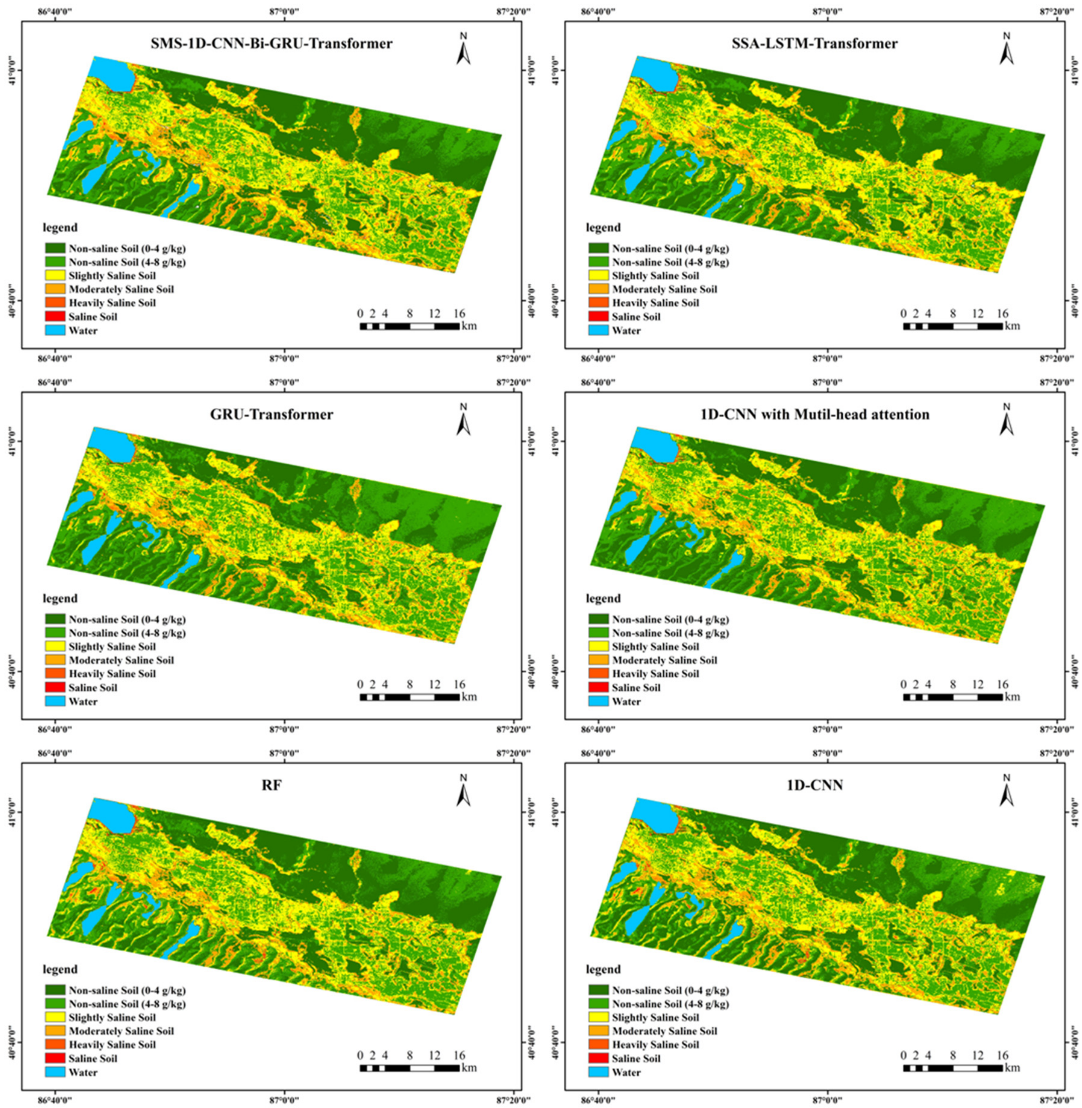


Figure 13. Visual Comparisons of Soil Salinity Retrievals Using Various Models.

5.5. Ablation Study: Unpacking Component Contributions

A structured ablation study (Table 6) was conducted to isolate the impact of each architectural component:

- Removing multi-scale extraction & scale attention caused the most pronounced decline in accuracy, particularly increasing RMSE to 0.972 $g \cdot kg^{-1}$. This confirms that multi-scale reasoning and dynamic feature weighting are essential for representing the heterogeneous spatial signatures of salinity.
- Removing the Bi-GRU-Transformer module led to a clear drop in R^2 (0.938) and increased MAE, underscoring the module's role in capturing sequential dependencies (via Bi-GRU) and global feature interactions (via Transformer) that are missed by pure CNN architectures.
- Removing only scale attention resulted in a moderate but consistent degradation across metrics, indicating that attention not only prioritizes informative scales but also serves as a bridge that optimally conditions feature for downstream sequence-global processing.
- The full model achieved the best balance, demonstrating that the components operate synergistically: multi-scale CNNs extract hierarchical features, scale attention reweights and primes them, and Bi-GRU-Transformer models their contextual and global relationships.

 doi:10.20944/preprints202604.1186.v1

These results validate the design rationale that no single mechanism is sufficient; rather, the interplay of scale-aware feature learning, importance weighting, and hybrid sequential-global modeling is key to high-fidelity salinity retrieval.

Table 6. Ablation Experimental Analysis.

Methods	R^2	RMSE/(g·kg-1)	MAE/(g·kg-1)	MAPE/%
1D-CNN with Bi-GRU-Transformer	0.930	0.972	0.619	6.704
MS-1D-CNN	0.922	1.062	0.676	7.011
SMS-1D-CNN with Scale attention	0.938	0.896	0.59	6.759
SMS-1D-CNN with Transformer	0.940	0.876	0.576	6.368
MS-1D-CNN with Bi-GRU-Transformer	0.943	0.884	0.547	6.519

5.6. Optimized Feature Extraction Strategy

Experiments on convolutional kernel scales (Table 7) revealed that employing four different kernel sizes (1, 3, 5, 7) yielded optimal performance (R^2 = 0.952, RMSE = 0.867 g·kg^{-1}). This configuration balances receptive field diversity with parameter efficiency. Smaller or larger scale sets led to either insufficient multi-scale coverage or feature redundancy, respectively, confirming that a deliberate multi-scale design is critical for modeling the varied spatial manifestations of soil salinity.

Table 7. Accuracy Evaluation of Different Convolutional Kernel Scales.

Scales Number	R^2	RMSE/(g·kg-1)	MAE/(g·kg-1)	MAPE/%
1	0.931	1.079	0.631	6.754
2	0.940	0.942	0.577	6.298
3	0.947	0.912	0.559	5.523
4	0.952	0.867	0.451	4.853
5	0.949	0.918	0.499	5.381

6. Conclusions

This study proposed a novel deep learning framework for soil salinity retrieval, integrating a scale-attention-optimized multi-scale 1D-CNN with a Bi-GRU-Transformer mechanism. The model was designed to address key limitations in existing methods, including inadequate multi-scale feature utilization, poor handling of sequential-spectral relationships, and insufficient global interaction modelling in multi-source remote sensing data.

Experimental results in the arid oasis region of Weili County, Xinjiang, demonstrated that the proposed model achieved a coefficient of determination (R^2) of 0.952 and a root mean square error (RMSE) of 0.867 g·kg^{-1}, representing an improvement of 2.8% in R^2 and 18.9% in RMSE over the best-performing baseline. The ablation study confirmed that each component—multi-scale feature extraction, scale attention, and Bi-GRU-Transformer—contributed synergistically to the model's performance. Specifically, the scale attention mechanism effectively prioritized discriminative features across spectral and spatial scales, while the Bi-GRU-Transformer module captured both local sequential patterns and global feature dependencies, enabling robust salinization mapping even under high spatial heterogeneity.

The model not only exhibited strong numerical accuracy but also produced spatially consistent and ecologically plausible salinity maps, correctly identifying salinity hotspots in low-lying and poorly drained areas. This underscores its practical utility for precision land management and salinity monitoring in vulnerable irrigated ecosystems.

In summary, this research contributes a scalable and interpretable deep learning architecture that advances the retrieval of soil salinity from multi-source remote sensing data. Future work will focus on extending the model to broader geographic regions, integrating temporal dynamics, and

 doi:10.20944/preprints202604.1186.v1

exploring its adaptability to other land degradation phenomena within a unified remote sensing analytics framework.

Author Contributions: All authors contributed to the study conception and design. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the High-performance GPU Server (TX321203) Computing Centre of the National Education Field Equipment Renewal and Renovation Loan Financial Subsidy Project of China University of Geosciences, Wuhan. This work was also supported by the National Key R&D Program "Integrated Ecological Restoration Standard System and Specification Research for Mountains, Rivers, Forests, Fields, Lakes, Grasslands, and Deserts" (2022YFF1303205), in part by the Opening Fund of Key Laboratory of Geological Survey and Evaluation of Ministry of Education under Grant No. GLAB2024ZR01.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Jiang, K.; Teuling, A.J.; Chen, X.; Huang, N.; Wang, J.; Zhang, Z.; Gao, R.; Men, J.; Zhang, Z.; Wu, Y.; Cai, L.; Huang, Z.; Ma, Z.; Pan, Z. Global land degradation hotspots based on multiple methods and indicators. Ecol. Indic. 2024, 158, 111462. https://doi.org/10.1016/j.ecolind.2023.111462.
2. Ondrasek, G.; Rengel, Z. Environmental salinization processes: Detection, implications & solutions. Sci. Total Environ. 2021, 754, 142432. https://doi.org/10.1016/j.scitotenv.2020.142432.
3. Li, Y.; Chang, C.; Wang, Z.; Zhao, G. Upscaling remote sensing inversion and dynamic monitoring of soil salinization in the Yellow River Delta, China. *Ecol. Indic.* 2023, *148*, 110087. https://doi.org/10.1016/j.ecolind.2023.110087.
4. Devkota, K.P.; Devkota, M.; Rezaei, M.; Oosterbaan, R. Managing salinity for sustainable agricultural production in salt-affected soils of irrigated drylands. *Agric. Syst.* 2022, *198*, 103390. https://doi.org/10.1016/j.agsy.2022.103390.
5. Zhang, Z.; Niu, B.; Li, X.; Kang, X.; Wan, H.; Shi, X.; Li, Q.; Xue, Y.; Hu, X. Inversion of soil salinity in China's Yellow River Delta using unmanned aerial vehicle multispectral technique. *Environ. Monit. Assess.* 2023, *195*, 245.
6. Du, X.; Zheng, X.; Lu, X.; Doudkin, A.A. Multisource remote sensing data classification with graph fusion network. *IEEE Trans. Geosci. Remote Sens.* 2021, *59*, 10062–10072. https://doi.org/10.1109/TGRS.2020.3047130.
7. Zhao, J.; Zhang, D.; Shi, B.; Zhou, Y.; Chen, J.; Yao, R.; Xue, Y. Multi-source collaborative enhanced for remote sensing images semantic segmentation. *Neurocomputing* 2022, *493*, 76–90. https://doi.org/10.1016/j.neucom.2022.04.045.
8. Wang, Y.; Fan, L.; Tao, R.; Zhang, L.; Zhao, W. Research on cropping intensity mapping of the Huai River Basin (China) based on multi-source remote sensing data fusion. *Environ. Sci. Pollut. Res.* 2022, *29*, 12661–12679.
9. Li, J.; Hong, D.; Gao, L.; Yao, J.; Zheng, K.; Zhang, B.; Chanussot, J. Deep learning in multimodal remote sensing data fusion: A comprehensive review. *Int. J. Appl. Earth Obs. Geoinf.* 2022, *112*, 102926. https://doi.org/10.1016/j.jag.2022.102926.
10. Wang, Z.; Chen, T.; Zhu, D.; Jia, K.; Plaza, A. RSEIFE: A new remote sensing ecological index for simulating the land surface eco-environment. *J. Environ. Manag.* 2023, *326*, 116851. https://doi.org/10.1016/j.jenvman.2022.116851.
11. Zhang, Z.; Fan, Y.; Zhang, A.; Jiao, Z. Baseline-based soil salinity index (BSSI): A novel remote sensing monitoring method of soil salinization. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2022, *16*, 202–214. https://doi.org/10.1109/JSTARS.2022.3223935.
12. Han, Y.; Ge, H.; Xu, Y.; Zhuang, L.; Wang, F.; Gu, Q. Estimating soil salinity using multiple spectral indexes and machine learning algorithm in Songnen Plain, China. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2023, *16*, 7041–7050. https://doi.org/10.1109/JSTARS.2023.3274579.

 doi:10.20944/preprints202604.1186.v1

13. Ma, S.; He, B.; Ge, X.; Luo, X. Spatial prediction of soil salinity based on the Google Earth Engine platform with multitemporal synthetic remote sensing images. *Ecol. Inform.* 2023, *75*, 102111. https://doi.org/10.1016/j.ecoinf.2023.102111.
14. Lotfollahi, L.; Delavar, M.A.; Biswas, A.; Fatehi, S.; Scholten, T. Spectral prediction of soil salinity and alkalinity indicators using visible, near-, and mid-infrared spectroscopy. *J. Environ. Manag.* 2023, *345*, 118854. https://doi.org/10.1016/j.jenvman.2023.118854.
15. Avdan, U.; Kaplan, G.; Matci, D.K.; Avdan, Z.Y.; Erdem, F.; Mizik, E.T.; Demirtas, I. Soil salinity prediction models constructed by different remote sensors. *Phys. Chem. Earth* 2022, *128*, 103230. https://doi.org/10.1016/j.pce.2022.103230.
16. Periasamy, S.; Ravi, K.P.; Tansey, K. Identification of saline landscapes from an integrated SVM approach from a novel 3-D classification schema using Sentinel-1 dual-polarized SAR data. *Remote Sens. Environ.* 2022, *279*, 113144. https://doi.org/10.1016/j.rse.2022.113144.
17. Wang, D.; Cao, W.; Zhang, F.; Li, Z.; Xu, S.; Wu, X. A review of deep learning in multiscale agricultural sensing. *Remote Sens.* 2022, *14*, 559. https://doi.org/10.3390/rs14030559.
18. Barbedo, J.G.A. A review on the combination of deep learning techniques with proximal hyperspectral images in agriculture. *Comput. Electron. Agric.* 2023, *210*, 107920. https://doi.org/10.1016/j.compag.2023.107920.
19. Das, A.; Bhattacharyya, B.K.; Setia, R.; Jayasree, G.; Das, B.S. A novel method for detecting soil salinity using AVIRIS-NG imaging spectroscopy and ensemble machine learning. *ISPRS J. Photogramm. Remote Sens.* 2023, *200*, 191–212. https://doi.org/10.1016/j.isprsjprs.2023.04.018.
20. Jianya, G.; Yansheng, L. Can quantitative remote sensing and machine learning be integrated? *Earth Sciences* 2022, *47*(10), 3911–3912.
21. Wang, N.; Peng, J.; Xue, J.; Zhang, X.; Huang, J.; Biswas, A.; He, Y.; Shi, Z. A framework for determining the total salt content of soil profiles using time-series Sentinel-2 images and a random forest-temporal convolution network. *Geoderma* 2022, *409*, 115656. https://doi.org/10.1016/j.geoderma.2021.115656.
22. Nguyen, H.D.; Van, C.P.; Nguyen, T.G.; Dang, D.K.; Pham, T.T.N.; Bui, Q.-T. Soil salinity prediction using hybrid machine learning and remote sensing in Ben Tre province on Vietnam's Mekong River Delta. *Environ. Sci. Pollut. Res.* 2023, *30*, 74340–74357.
23. Chen, Y.; Liu, L.; Chen, X.; Wei, Z.; Sun, X.; Yuan, C.; Gao, Z. Data driven three-dimensional temperature and salinity anomaly reconstruction of the northwest Pacific Ocean. *Front. Mar. Sci.* 2023, *10*, 1121334. https://doi.org/10.3389/fmars.2023.1121334.
24. Garajeh, M.K.; Blaschke, T.; Haghi, V.H.; Weng, Q.; Kamran, K.V.; Li, Z. A comparison between Sentinel-2 and Landsat 8 OLI satellite images for soil salinity distribution mapping using a deep learning convolutional neural network. *Can. J. Remote Sens.* 2022, *48*, 452–468. https://doi.org/10.1080/07038992.2022.2056435.
25. Zheng, W.; Zheng, K.; Gao, L.; Zhangzhong, L.; Lan, R.; Xu, L.; Yu, J. GRU–Transformer: A novel hybrid model for predicting soil moisture content in root zones. *Agronomy* 2024, *14*, 432. https://doi.org/10.3390/agronomy14030432.
26. Li, Z.; Yuan, Q.; Zhang, L. Geo-intelligent retrieval framework based on machine learning in the cloud environment: A case study of soil moisture retrieval. *IEEE Trans. Geosci. Remote Sens.* 2023, *61*, 4502615. https://doi.org/10.1109/TGRS.2023.3280591.
27. Nabi, M.M.; Senyurek, V.; Lei, F.; Kurum, M.; Gurbuz, A.C. Quasi-global assessment of deep learning-based CYGNSS soil moisture retrieval. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2023, *16*, 5629–5644. https://doi.org/10.1109/JSTARS.2023.3287591.
28. Yao, J.; Wang, Z.; Oliveira, R.F.d.; Wang, Z.-Y.; Huang, L. A deep learning method for the long-term prediction of plant electrical signals under salt stress to identify salt tolerance. *Comput. Electron. Agric.* 2021, *190*, 106435. https://doi.org/10.1016/j.compag.2021.106435.
29. Zhang, S.; Tian, J.; Lu, X.; Tian, Q.; He, S.; Lin, Y.; Li, S.; Zheng, W.; Wen, T.; Mu, X.; Zhang, J.; Li, Y. Monitoring of chlorophyll content in local saltwort species *Suaeda salsa* under water and salt stress based on the PROSAIL-D model in coastal wetland. *Remote Sens. Environ.* 2024, *306*, 114117. https://doi.org/10.1016/j.rse.2024.114117.

30. Ivanda, A.; Šerić, L.; Žagar, D.; Oštir, K. An application of 1D convolution and deep learning to remote sensing modelling of Secchi depth in the northern Adriatic Sea. *Big Earth Data* 2024, *8*, 82–114. https://doi.org/10.1080/20964471.2023.2273058.
31. Ghasempour, R.; Aalami, M.T.; Saghebian, S.M.; Kirca, V.S.O. Analysis of spatiotemporal variations of drought and soil salinity via integrated multiscale and remote sensing-based techniques (Case study: Urmia Lake basin). *Ecol. Inform.* 2024, *81*, 102560. https://doi.org/10.1016/j.ecoinf.2024.102560.
32. Shi, J.; Wang, Y.; Yu, Z.; Li, G.; Hong, X.; Wang, F. Exploiting multi-scale parallel self-attention and local variation via dual-branch Transformer-CNN structure for face super-resolution. *IEEE Trans. Multimed.* 2023, *26*, 2608–2620. https://doi.org/10.1109/TMM.2023.3301225.
33. Zhang, Y.; Shi, J.; Wang, J.; Zong, Y.; Zheng, W.; Zhao, G. MaskFusionNet: A dual-stream fusion model with masked pre-training mechanism for rPPG measurement. *IEEE Trans. Circuits Syst. Video Technol.* 2024, *34*, 11521–11534. https://doi.org/10.1109/TCSVT.2024.3422849.
34. Wang, J.; Li, C.; Li, L.; Huang, Z.; Wang, C.; Zhang, H.; Zhang, Z. InSAR time-series deformation forecasting surrounding Salt Lake using deep transformer models. *Sci. Total Environ.* 2023, *858*, 159744. https://doi.org/10.1016/j.scitotenv.2022.159744.
35. Gad, M.M.; Mohamed, M.R. Estimating salinity using remote sensing data. *Journal of Al-Azhar University Engineering Sector.* 2022, *17*(65), 1143–1156.
36. Cheraghi, M.; Motesharezadeh, B.; Mousavi, S.M.; Ma, Q.; Ahmadabadi, Z. Silicon (Si): A regulator nutrient for optimum growth of wheat under salinity and drought stresses—A review. *J. Plant Growth Regul.* 2023, *42*, 5354–5378.
37. Abbas, J.A.A. Soil salinity assessment by using spectral salinity indices in Al-Sweira project middle of the Iraqi alluvial plain. *Int. J. Environ. Sci. Technol.* 2023, *20*, 10847–10860.
38. Cui, X.; Han, W.; Zhang, H.; Cui, J.; Ma, W.; Zhang, L.; Li, G. Estimating soil salinity under sunflower cover in the Hetao Irrigation District based on unmanned aerial vehicle remote sensing. *Land Degrad. Dev.* 2023, *34*(1), 84–97. https://doi.org/10.1002/ldr.4445.
39. Venâncio, J.B.; Dias, N.S.; Medeiros, J.F.; Morais, P.L.D.; Nascimento, C.W.A.d.; Sousa Neto, O.N.d.; Sá, F.V.d.S. Yield and morphophysiology of onion grown under salinity and fertilization with silicon. *Sci. Hortic.* 2022, *301*, 111095. https://doi.org/10.1016/j.scienta.2022.111095.
40. Aksoy, S.; Yildirim, A.; Gorji, T.; Hamzehpour, N.; Tanik, A.; Sertel, E. Assessing the performance of machine learning algorithms for soil salinity mapping in Google Earth Engine platform using Sentinel-2A and Landsat-8 OLI data. *Adv. Space Res.* 2022, *69*, 1072–1086. https://doi.org/10.1016/j.asr.2021.10.024.
41. Zhao, W.; Ma, H.; Zhou, C.; Zhou, C.; Li, Z. Soil salinity inversion model based on BPNN optimization algorithm for UAV multispectral remote sensing. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2023, *16*, 6038–6047. https://doi.org/10.1109/JSTARS.2023.3284019.
42. Kumar, G.A.; Katiyar, A.; Srinivasan, K. Elevating IDS capabilities: The convergence of SVM, deep learning, and RFECV in network security. In *Proceedings of the 2024 Second International Conference on Emerging Trends in Information Technology and Engineering (ICETITE)*, Vellore, India, 22–23 February 2024. https://doi.org/10.1109/ic-ETITE58242.2024.10493239.
43. Wang, Z.; Xia, H.; Zhang, J.; Yang, B.; Yin, W. Imbalanced sample fault diagnosis method for rotating machinery in nuclear power plants based on deep convolutional conditional generative adversarial network. *Nucl. Eng. Technol.* 2023, *55*, 2096–2106. https://doi.org/10.1016/j.net.2023.02.036.
44. Wang, Y.; Xie, M.; Hu, B.; Jiang, Q.; Shi, Z.; He, Y.; Peng, J. Desert soil salinity inversion models based on field in situ spectroscopy in Southern Xinjiang, China. *Remote Sens.* 2022, *14*(19), 4962. https://doi.org/10.3390/rs14194962.
45. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention is all you need. In *Advances in Neural Information Processing Systems 30 (NeurIPS 2017)*; 2017; pp. 5998–6008.
46. Schippers, H.; Schüler, C.; Sliwa, B.; Wietfeld, C. System modeling and performance evaluation of predictive QoS for future tele-operated driving. In *Proceedings of the 2022 IEEE International Systems Conference (SysCon)*, 25–28 April 2022. https://doi.org/10.1109/SysCon53536.2022.9773810.

 doi:10.20944/preprints202604.1186.v1

47. Shi, H.; Hellwich, O.; Luo, G.; Chen, C.; He, H.; Ochege, F.U. A global meta-analysis of soil salinity prediction integrating satellite remote sensing, soil sampling, and machine learning. *IEEE Trans. Geosci. Remote Sens.* 2022, *60*, 4505815. https://doi.org/10.1109/TGRS.2021.3109819.
48. Chen, D.; Yongchareon, S.; Lai, E.M.-K.; Yu, J.; Sheng, Q.Z.; Li, Y. Transformer with bidirectional GRU for nonintrusive, sensor-based activity recognition in a multiresident environment. *IEEE Internet Things J.* 2022, *9*, 23716–23727. https://doi.org/10.1109/JIOT.2022.3190307.
49. M., S.; C.D., J.; Lingappa, M. 1D convolutional neural networks-based soil fertility classification and fertilizer prescription. *Ecol. Inform.* 2023, *78*, 102295. https://doi.org/10.1016/j.ecoinf.2023.102295.
50. Guo, F.; Wang, P.; Tansey, K.; Zhang, Y.; Li, M.; Liu, J.; Zhang, S. A novel transformer-based neural network under model interpretability for improving wheat yield estimation using remotely sensed multi-variables. *Comput. Electron. Agric.* 2024, *223*, 109111. https://doi.org/10.1016/j.compag.2024.109111.
51. Victor, B.; Nibali, A.; He, Z. A systematic review of the use of deep learning in satellite imagery for agriculture. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2024, *18*, 2297–2316. https://doi.org/10.1109/JSTARS.2024.3501216.